1. introduction

Discuss different combination of : features to make the best of data.

models

2. Features

manual (15 not including id and label)

raw data

feature selection (maximum 15)

normalized

PCA (15)

(hidden layer of neural network)

3. Models

decision tree

knn

naive bayes

random forest

deep learning (“2 layer with 50 neurons” and “4 layers with 25 neurons”)

4. Experimental setup

Corpus

Prepocess

5. results of experiments

about model

Is DNN always best (DNN vs all)

The ability of DNN

IS emsemble learining better (Decision tree vs Random forest)

about feature

Is humen efforts worth or not (manunal vs feature selection)

What kind of feature is always best (all)

5. conclusions